

```

/clone="IMAGE:3659630"
/clone_lib="NCI_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/notes="Organ: Lung; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; transgenic model MMT-1, expression driven by
MTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by life technologies.
Investigator providing samples: Gilbert Smith, NIH"

```

```

BASE COUNT      108 a      174 c      129 g      111 t
ORIGIN

```

```

Query Match      18.1%; Score 517.2; DB 139; Length 522;
Best Local Similarity 99.4%; Pred. No. 1.3e-132;
Matches 519; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

OY 81 ggcctgtgtgagctccccctctctctgagcccgccgagggcgccctcgcctc 140
1 ggcctgtgtgagctccccctctctctgagcccgccgagggcgccctcgcctc 60
OY 141 cgcacactcgcctctctctctctctgagccctcttctcctcgcctcctc 200
61 cgcacactcgcctctctctctctctgagccctcttctcctcgcctcctc 120
OY 201 tcttcaactctctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 260
121 tcttcaactctctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
OY 261 ccggaagatagctcttaccatgagctcaatgacatgctgtgtgaagtgaaa 320
181 ccggaagatagctcttaccatgagctcaatgacatgctgtgtgaagtgaaa 240
OY 321 ccaatgtataagccgctgagccctcctcctcctcctcctcctcctcctc 380
241 ccaatgtataagccgctgagccctcctcctcctcctcctcctcctcctc 300
OY 381 cgtgagagctgcctcctcctcctcctcctcctcctcctcctcctcctcct 440
301 cgtgagagctgcctcctcctcctcctcctcctcctcctcctcctcctcct 360
OY 441 caagtgtagctctcctcctcctcctcctcctcctcctcctcctcctcct 500
361 caagtgtagctctcctcctcctcctcctcctcctcctcctcctcctcct 420
OY 501 gtagaacaagatgcctcctcctcctcctcctcctcctcctcctcctcct 560
421 gtagaacaagatgcctcctcctcctcctcctcctcctcctcctcctcct 480
OY 561 aggttgagtgaaatgagagagagagagagagagagagagagagagagag 602
481 aggttgagtgaaatgagagagagagagagagagagagagagagagagag 522

```

```

RESULT 10
LOCUS      AA122533      522 bp      mRNA      EST      17-FEB-1997
DEFINITION mp95c04.r1 Soares_thymus_2NDMT Mus musculus cDNA clone IMAGE:576966
5', mRNA sequence.
ACCESSION  AA122533
VERSION     AA122533.1 GI:1681596
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus.
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 522)
Marras, M., Hillier, L., Allen, M., Bowles, N., Dietrich, N., Dubuque, T.,
Gaisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Thelsting, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterson, R.
TITLE       The WashU-RHMI Mouse EST Project

```

```

JOURNAL      Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
              WashU-RHMI Mouse EST Project
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: mouseest@atson.wustl.edu
              This clone is available royalty-free through LIND; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              MGI:351614
              Seq primer: -28M13 rev2 from Amersham
              High quality sequence stop: 448.

```

```

FEATURES
source

```

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:576966"
/clone_lib="Soares_thymus_2NDMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/notes="vector: pT73D-pac (Pharmacia) with a modified
polylinker; Site:1: Not I; Site:2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer (5'
TGTTACCAATCTGAGTGGAGGCGCGCGGCTTTTCTTTTCTTTTCTTTTCTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."

```

```

Query Match      17.9%; Score 511; DB 2; Length 522;
Best Local Similarity 99.8%; Pred. No. 6.9e-131;
Matches 522; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

OY 2248 tgatacaaacataattcattcattcattcattcattcattcattcattc 2307
1 tgatacaaacataattcattcattcattcattcattcattcattcattcattc 60
OY 2308 agaaactgtcagctatgtaactgtaactgtaactgtaactgtaactgta 2367
61 agaaactgtcagctatgtaactgtaactgtaactgtaactgtaactgtaact 120
OY 2368 tcttgctgtgtaagagcttctgtaattacacacagaaatcctgtgactcc 2427
121 tcttgctgtgtaagagcttctgtaattacacacagaaatcctgtgactcc 179
OY 2428 ataatgtgacagcttgaagacttcaaaccttgaatcacaagaatagatgct 2487
180 ataatgtgacagcttgaagacttcaaaccttgaatcacaagaatagatgct 239
OY 2488 ctattcttaagagagagactttaaagaagccctgagcttccatataagaatt 2547
240 ctattcttaagagagagactttaaagaagccctgagcttccatataagaatt 299
OY 2548 gtcaaaagagagcttcttgaagagagagagagagagagagagagagagag 2607
300 gtcaaaagagagcttcttgaagagagagagagagagagagagagagagag 359
OY 2608 gtagcccaagaagactacagagctgtgtccctgaagagagagagagagagag 2667
360 gtagcccaagaagactacagagctgtgtccctgaagagagagagagagagag 419
OY 2668 tctaaagtcggggaagagtgacagagcttccctgtgtgtgtgtgtgtgtgt 2727
420 tctaaagtcggggaagagtgacagagcttccctgtgtgtgtgtgtgtgtgt 479

```

OY 2728 agctggcctggggaagacattctgctcctagtcagaccctg 2770
 |||||
 Db 480 AGCTGGCTGGGGAAGACATGTGGCTCTAGTCCAGCCCTG 522
 RESULT 11
 BE894295 752 bp mRNA EST 29-SEP-2000
 LOCUS 601437666f1 NIH_MGC_72 Homo sapiens CDNA clone IMAGE:3922529 5',
 DEFINITION mRNA sequence.
 ACCESSION BE894295
 VERSION BE894295.1 GI:10356518
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 752)
 NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/
 Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: ATCC/DCTD/DRP
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LDCM744 row: d column: 18
 High quality sequence stop: 654.
 FEATURES
 SOURCE
 1..752
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3922529"
 /clone_lib="NIH_MGC_72"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site:1; Notif.
 Site:2; Salt: Cloned unidirectionally. Primer: oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 BASE COUNT 223 a 194 c 192 g 143 t
 Query Match 17.88; Score 508.4; DB 137; Length 752;
 Best Local Similarity 84.84; Pred. No. 4,1e-130;
 Matches 616; Conservative 0; Mismatches 106; Indels 4; Gaps 4;
 OY 892 attatggccaaggatgaatccatctagatgacacagtcacagcagcagcaaaaaag 951
 |||||
 Db 1 AATATGGCCAGGGGATCAATCCGATTAGCCCACTGCCCCAGATCCAGAGCAAAAAAG 60
 OY 952 aagaagagcagaagatcatcgtctccctacagaagcaggtcttcacagtcgagaggttg 1011
 |||||
 Db 61 AGAAGAGGACGAGATGACACGCTCTCACAGAGCGAGCGCTCCCGCCGAGGAGAGTTG 120
 OY 1012 tgatcgaggttaaggtgtggcatcacacatcagcagcagagagtggtgtaccataaaggttg 1071
 |||||
 Db 121 TGAATGAGGTGAAAGGTTGGAACCACTGACAGAAAGGAACGGACCAACAAAGAGTGG 180
 OY 1072 ccaagcgtaactgctgtgagacacatcgtcgagatccgtgggttcacaaagtcccccagcgc 1131
 |||||
 Db 181 CCAAGGCAATGACGCGAGAACATGCTGAGATCTTGTTCAAAAGTCCCGCAGCGC 240
 OY 1132 agctcgccaagcagcagcacaatcagagagaagaaactcccaataaagaaacagagagcg 1191
 |||||
 Db 241 AGCCCAACCAAAACCCGACTCAAGTCAGAGAGAACACCCCATAAAGAACCAAGG -GATG 299

OY 1192 gaaggaagtaacgtttcttgaaacctagccctggggatgtaaaatggaactagtaacagag 1251
 |||||
 Db 300 GAAGAAAAGTAACCTTTTGTGAACCTGCTGGGGATGAAAGGCACTAGTAATAAG 359
 OY 1252 acgaagagttcagatgctcttcttagccatcagcagcagcagcagcagcagcagcagc 1311
 |||||
 Db 360 AGGATGAGTTGAGATGCTTCTTCTAGTCATCAGCAGCAGCGCGCGGAAATTCCTCCCA 419
 OY 1312 tggcgaggaagttgcccagcgtgctggggttagtaaggaagacaccacacaaagattcca 1371
 |||||
 Db 420 TGGTCCCGGAGGTGGCCAGCGCTGTAGAGATTAGTCAAGGACATCACACCAAAATTTTA 479
 OY 1372 ccaaggaagcctcacaatctctgccaagaggaaggttaactgccaatagatagccagagttgt 1431
 |||||
 Db 480 CCAAGACAGCTCCGAAATCTGCGAAGCCAGCGGTACCTGCTGCTGCTGCTGCTGCTGCTG 539
 OY 1432 tgacggggagcctcggcccaagcagcagcagcagcagcagcagcagcagcagcagcagc 1490
 |||||
 Db 540 AGTATGGGGGCGACCTCGCCACAGCGAGATCATTTTAAAGAAATACATCTCTCAGCG 599
 OY 1491 caagtaaccacatgagcctcgcacatagaccctcgcagcagcagcagcagcagcagcagc 1550
 |||||
 Db 600 CACGTACCCCATGCA -CTCTCAGAGATCTCTGAGCA -TGGCTATCTTCCAGACTC 657
 OY 1551 caggatccaggttgatatacaagaatttcccaagacacacagaacagagtgatctc 1610
 |||||
 Db 658 CAGGATTCGGGTTGATCAAAAGACTTCCCAAAAAAAGCAAGAAATTTGTTCTCTATCA 717
 OY 1611 ctcac 1616
 |||||
 Db 718 ATGCTC 723
 RESULT 12
 AM542549/c 511 bp mRNA EST 31-AUG-2000
 LOCUS C0158H10-3 Mouse E7.5 Extraembryonic Portion CDNA Library Mus
 DEFINITION musculus CDNA clone C0158H10 3', mRNA sequence.
 ACCESSION AM542549
 VERSION AM542549.1 GI:7184966
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 511)
 Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grabovac
 ,M.J., Pantano,S., Sano,Y., Piao,Y., Nagataja,R., Dol,H., Wood,W.H.
 III, Becker,K.G. and Ko,M.S.H.
 Genome-wide expression profiling of mid-gestation placenta and
 embryo using a 15,000 mouse developmental CDNA microarray
 Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@sun.gsc.nia.nih.gov
 Plate: C0158 row: H column: 10
 Seq primer: -21M3 Forward
 High quality sequence stop: 511
 POLY-A=yes.
 FEATURES
 SOURCE
 1..511
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="C0158H10"
 /clone_lib="Mouse E7.5 Extraembryonic Portion CDNA
 Library"
 /sex="unknown"
 /dev_stage="7.5dpc Embryo"

396 396
38383 TCTTCTCTGAAAGCTATTTCTGGGATTGCTGCTGCTGAAATACAGCA 38334
396 396
38333 GGAGCCCAAGCAGTAGTCCCGCATGCAAAAGTCTTTATCATTCCTCG 38284
396 396
38283 ATTCCTCGCGCTCCATGCGTCCAGGAGGAGGCGCAGCTGGGATTGA 38234
396 396
38233 TGGTCTCATGCGTTTTCAGGCTGGATGATGCTGCTTACCGT 38184
38183 Ser. AsnlysgluaspqluphearmetProtyrleu 408
409 SerHISGLINGluProAlaGlyLeuLeuProMetValProgluValAl 425
38133 AGTATGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 38084
425 AGlnAlaValAlValSerGlnGlyHISThrLysAspPheThrArga 442
38083 CCAGCTGCTAGAGTACTGCAAGGACATCACCAAGATTATTAACAGGG 38034
442 laAlaProAsnProAlaLysAlaThrValThrAlaMetIleAlaArgGlu 458
38033 CAGCTCGGAATCCGTCAGGAGGCGGATGCTGCTGCTGCTGCTGCTG 37984
459 leuLeuTyrglyGlyThSerProThrAlaGluThrIleLeuLysAsnAs 475
37983 TTGTTGATGGGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37934
475 nileSerSerGlyHISValProHISGlyProLeuThrArgProSerGlu 492
37933 CATCTCTTCAGGCGCAGTACCCTGCTGCTGCTGCTGCTGCTGCTGCT 37884
492 InLeuAspTyrlLeuSerArgValGlnGlyPheGln..... 503
37883 AACTGACATATCTTCCAGAGTCCAGGATTCAGGTAATCTGTGGCT 37834
37833 GAGCTGTGATGGCTGTGCTGCTGAGACAGTCACTTTTGGCTTCTCAGA 37784
503 503
37783 AAATGAACCTTGAACCTTGTACTTTTCTTAATCAAGTCTGTCCAGG 37734
503 503
37733 GATGGCTGGGAGCACTGATGAAGGCTAGAGGTATAGAGTATTGTA 37684
503 503
37683 AGCATATGTGCTAAGAAATATCTTATCTTGTGCTAGTAAGATTCAA 37634
503 503
37633 CATGCAAGCCAAACCTAAAGCTTCTAACCACATAGAGTGTGTAAT 37584
503 503
37583 TGATATTCAAGAACAGACTCATCTGCTGACACTTGTCTGTGTTAGGA 37534
503 503
37533 AAGGATATGCTTCTCCAGTCTTGGCTTCTGAGACAGACACTAGTTC 37484

503 503
37483 CTGGAAGCTTCTGCGCTGACAGGCGAGTGTATGATGCGCAGTGGGG 37434
503 503
37433 CTGAGTTGGAAAAAGAGCTGTGATCTGCTTGTGGGGCTGAGAGTGC 37384
503 503
37383 AGCCCTCCAGCCCACTCACTGCTGCTTCCCTACCTAGCTCAGCTTTC 37334
504 ValGluTyrlLysAspPheProLysAsnAsnlyAsng 516
37333 TTTCCCGCTTAGTGAATACAAAGACTTCCCAAAACAAAGAACG 37284
516 IupheValSerLeuIleAsnGysSerSerGlnProProLeuIleSerHis 532
37283 AATTGTATCTCTTATCAATTTGCTCTGCTGAGCCACCTCTGATCAGCAT 37234
533 GlyIleGlyLysAspValGluSerCysHISAspMet 544
37233 GGTATCGGCAAGAGATGTGAGTCTGCTGCTGCTGCTGCTGCTGCTG 37198
seq_name: gb_in3:DROSTAUFE
seq_documentation_block: 5237 bp mRNA INV 26-APR-1993
LOCUS DROSTAUFE 5237 bp mRNA complete cds.
DEFINITION D.melanogaster mRNA, complete cds.
ACCESSION M69111
VERSION M69111.1 GI:156505
KEYWORDS
SOURCE Drosophila melanogaster 0-4 hour embryo cDNA to mRNA.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 5237)
AUTHORS Johnston,D.S., Beuchle,D. and Nusslein-Volhard,C.
TITLE Staufen, a gene required to localize maternal RNAs in the
Drosophila egg
JOURNAL Cell 66, 51-53 (1991)
MEDLINE 91300552
FEATURES
Source location/Qualifiers
1..5237
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/dev_stage="0-4 hour embryo"
1..5237
275..3355
/codon_start=1
/protein_id="AA173062.1"
/db_xref="GI:156506"
/translation="MOHNVHAARPAHPIRAHHSHSHAHNHLPMEOHILPSLQOO
OOPPPPOOPPHRDHLRLNHHHLHAOQOQOQSSNOAAVAAGAAVHHGNNINS
GSNISNSNMOKIROOHNLSSNGILGNOPPPPOAPFNPALVAVNOLPH
PPHMAHILGSLAAPPYYSOAKPAKYNHNGSNNSNSNSNAPKALIONT
YRNOKVYVPPVOEATPVPEPVYTNATNSTNSNYLASEPVOTSOKEPTROE
PASADHYSTGINDATGALSNEDESSSGRGKOTTPHCLVIELARKYKITHQRLTE
KAPAHKTFVTMLDDEESADGFKIKKOHCLASRAIETMKRHPKTRSEGG
PMRHTTPTVELNALAMKLSQRTFYLLDPQIIPEDSVIPEFAGGLTAPGPMQ
PPPPAYALRORLNGFEPVIPSOPMHHPFHGPOPPPKPSRFALPPLGAVHH
GPNPPSPVPTPPSKITFVGKOKFVGIGRTLOAKHDAARALQVLTQISASEA
LEDSNDEGKSPISOVHEIGIKRNTVHKVLEEDPAHKKNTIACIVGSIYGE
GNGKVSRRRAAKMVELQKLPETTKOTPLKRIKIVKPGKSGAAREGSVSGTD
GTMOTGKRRRLNPPKDKLIMDDADNPITKLIQOQTRKEPIFEIITAKNET
ARRREFYVEVSGSTARCTGNSKKLAKRMAALFELEAVOYTPRNETDSSECT
SATNSAVTAPAVEATAEGKVPVAVATVPVPGILILEQNKRPARKRQIVTKSVES
KEEBANKVAVAAEENSNSNSNSNSGSDQATEAASESLNSTGNTSGVS
NSSNVGANTDGNNAHSEKNTESSNSTSTQAGVHMKQLLYLSKLDFEVNFSY
PKGHNRELTIVLSTHPQICGVGKSSSESDNDAASNLKILSKLGNNAK"

BASE COUNT 1596 a 1291 c 1204 g 1146 t
Origin Chromosome 55A, right arm.

alignment_scores: Length: 644
 Quality: 716.00 Gaps: 18
 Ratio: 2.081 Percent Identity: 32.143
 Percent Similarity: 53.416

Alignment_block:
 US-09-316-048-6 x DROSTAFUFEN ..

Align seg 1/1 to: DROSTAFUFEN from: 1 to: 5237

```

71 11ethrProthrValGluLeuAsnAlaLeuLysMetLysLeuGlyLys 87
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1466 ATTACGCCAGCGGTGAATGAATGAACTGCGCATGAGTTGGGACAGCG 1515
   87 sPromeTyr.....LysProValAsp..... 94
   : |||||
1516 CACCTCTACCTGCTGATCCCATGATTCCTCCACGCGATTCATAG 1565
   94 ..... 94
1566 TGCCCGCGAGTTTGGCGGCTCATTTACTAACTGCACGACGACGAGG 1615
   95 ...ProTyrSerArgMetGlnSerThrTyrAsnTyrAsnMetArgGly 110
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1616 ATGCGGCAACCACTCCACCCTGCTTACGCTCGCGCAGCGGCTTGG 1665
   110 Y..... 110
1666 CAATGATTTTGCCCATTCATCGCAGCGGATGCATCCGATTTCTTC 1715
   111 .....AlaTyrProTyrArgTyrPheTyrProPhePro 121
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1716 ATGACCTGCTGACGCGACCGTTTCTCGAAGTTTCCATCAGCGTTTGG 1765
   122 ValProLeuLeu.....TyrGlnValGluLeuSerValGlyGly 126
   : ||||| |||||
1766 CTGCCACCGCCCTCGAGCTCATGATCATGACGACGAGTCCCATTT 1815
   127 ..... 127
1816 TCCATCCCTGCTACGCCACCCAGCAAAATCAACCTGTTTGGGCAAGC 1865
   136 1ngInPheAsnGlyLysGlyLysThrArgGlnAlaAlaLysHisAspAla 152
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1866 AGAATTTTGTGGCATTCGACGACCACTGCGCAAGCAAGCAATGACGCT 1915
   153 AlaAlaLysAlaLeuArgLysLeuGlnAsnGluProLeu.....Pr 166
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1916 GCGGGAAGAGGCTGACAGTTCTCAAGACACAGGCCATATCTGCTCTGA 1965
   166 OGluArgLeuValAlaGlnGlyArgGluSerGluGluGluAsnLeuAsn 183
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1966 GGAGGCTCTCAG.....GATTCGATGATGAGGCTGACACAGA 2003
   183 YssSerGluLysSerGlnValPheGluIleAlaLeuLysArgAsnLeuPro 199
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2004 AGTGGCCCATTCGACAGTTCACGAGATCGAATTAAGCGCAACATGACG 2053
   200 ValAsnPheGluValAlaArgGluSerGlyProProHisMetLysAsnPh 216
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2054 GTTCATTTTAAGTGTGCGGAGAGGCGCGGCGGACATGAAAGCTT 2103
   216 eValThrLysValSerValGlyGluPheValGlyGluGlyLys 233
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2104 TATTAACGCTGCTATTTGTTGATGATGATTCACCAAGAGAAAGAAAG 2153
   233 eTyrLysLysSerLysLysAsnAlaAlaIleAlaValLeuGluGluLeu 249
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2154 GCAAAAAAGTGTCAAGAGAGCTGCGGGAAGAAAGATGCTGCTGAGTTG 2203
   250 LysLysLeuProProLeu...ProAlaValGluArgValLysProArgTl 265

```

```

2204 CAAAGCTGCGCCCTTACGCCACCAAGCAGCGCATTTGAAGCGGAT 2253
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 eLysLysLysThr..... 269
   ||||| |||||
2254 TAAGGTAAAGACACAGCAAGCAAAAGTGTGACGTGCGCCGAGAGTTGCG 2303
   270 .....LysProIleValLys 274
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2304 TTGTGTCGGCAGCATGATGATCTATATGCAAAAGGGAAGCGGAGAGAGCT 2353
   275 ProGlnThrSerPro.....GluTyrGlyGlnGlyI 285
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2354 AAAGCCCTCAATCCGCCGGAAGCAAGCTGATGACATGGACGATGGGGA 2403
   285 eAsnProIleSerArgLeuAlaGlnIleGlnAlaLysLysGluLysG 302
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2404 TAATCCCATTAACAAGCTAACTCACTGACGACAGCGGTAAAGAAAGG 2453
   302 LuProGluTyrThrLeuLeuThrGluArgGly.....LeuProArgArg 316
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2454 AACCCATTTTCAGCTGATAGCCAAAACGCAATGAAACCGCTCGACGA 2503
   317 ArgGluPheValMetGlnValLysValGlyAsnHisThrAlaGluGlyTh 333
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2504 CGGGAATTCGTTATGAGGTCTCCGCGGAGGAGCAGCCCGTGGAAAC 2553
   333 rGlyThrAsnLysLysValAlaLysArgAsnAlaAlaGluAsnMetLeu 350
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2554 AGGCAACAGCAAGAAATTGGCAAGGAATGCTGCACAGGCTCTATTGG 2603
   350 LuIleLeuGlyPheLysValProGlnThrGlnProThrLysProAlaLeu 366
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2604 AACGTCTG.....GAAGCTGTCAAGTCACACCAACCAAGAAACAG 2647
   367 LysSerGluGlu..... 370
   : ||||| |||||
2648 TCATCGGAGGATGCTGCATTCGGCAACAATGTCGCCGCTTACAGCGC 2697
   371 .....LysThrProI 374
   ||||| |||||
2698 CGCAGTGAGCAACTGCGGAAGCAAAAGTGCCATGTGATGACCAACCCG 2747
   374 Le.....LysLysProGly 378
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2748 TGGTCCCATGCCAGCGCATTTAATCTACGCCAGAACAAAGAACCAAGCA 2797
   379 AspGlyArg...LysValThrPhePheGluProGlySerGlyAspGluAs 394
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2798 AAGAAAGAGGATCAAAATTGTAATGTCAAGTCAAGTGAATCCAAAGA 2847
   394 nGlyThrSerAsnLysGluAspGluPheArgMetProTyrLeuSerHisG 411
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2848 GGAAGAGGCCAATAAGAGGTGGCAGTTGCAAGTGAAGAAATGACAAACA 2897
   411 1ngInLeuProAlaGlyIleuProMetValProGluValAlaGlnAla 427
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2898 ACAGTGCAAACTCTGCG.....GACAGCAGCAACAGCAACG 2932
   428 ValGlyValSerGlnGlyHisThrLysAspPheThrArgAlaAlaArg 444
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2933 AGCGCGGATCGCAAGCC.....ACAGAGCGCGCTAG 2964
   444 OAsnProAlaLysAlaThrValThrAlaMetIleAlaArgGluLeuLeu 461
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2965 TAGAGCTGCACTGAATGCTCAACCGGACGCAATAG..... 3001
   461 rGlyGlyThrSerProThrAlaGluThrIleLeuLysAsnAsnIleSer 477
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3002 ..AGTGTGTGATAGCAACAGCAATGTTGAGCCAAACAGGATGTG 3049
   478 SerGlyHisValProHisGlyProLeuThrArgProSer..... 490
   : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

3050 AACNACCATGCCGAGACTAAAAACACAGAGAGTAGCAGCAGACTAC 3099
491 .....GluGlnLeuAspTyrL 496
3100 GAGTAACACACAAAGCGCTGAGTGCACATGAGGAGCAGCTTGTACC 3149
496 euSerArgValGlnGlyPheGlnValGluTyrLysAspPheProLysAsn 512
3150 TTAGTAAACTCCTCGACTTGTAGGTCAACTTCTCGACTATCCGAAAGGC 3199
513 AsnLysAsnGluPheValSerLeuIleAsnCysSerSerGlnProProLe 529
3200 AATCACAAGAGATTCTCTGACCATCGTGCATTTGTCACACATCCGCCGCA 3249
529 uIleSerHisGlyIleGlyLysAspValGluSerCysHisAspMetAlaA 546
3250 GATTGCCCATGGCGTGGCCAGAGCTCCGAGAGAGTCCGCAATGATGCTG 3299
546 IaleuAsnIleLeuLysLeuLeuSerGluLeu 556
3300 CAAGCAATGCTTGAAATCTTCAGCAAGCTG 3331

```



```

Db 60502 ACATCTCTTCTAGTGAAGTAAAGAACATTTGATGTGATGCAATCATGTGACTCT 60443
Oy 2946 ggtcagcagcgaacagtcctggaact-tcnaaaaccttgaaagtcctccacagctgta 3004
Db 60442 GGTTCAGCTGAAAAAGCCCTGAGCTTCTCAGAAACCTTGATCAATTTCCACTATGTGA 60383
Oy 3005 taattggacacatttagaattttaacttttagatgtaactgtaacttttcttcttcttcc 3064
Db 60382 CAATTTGGCCATTTTGAAGAAATTTCAAACTTCA---ATCATTTGGTCTCTTTTAAATTTT 60327
Oy 3065 atttctattcttgtaatagc-aacagagacttaaatgaactggaactcctgcttctttaaag 3122
Db 60326 ATTTTATTTTGGTTGTGTATATGAGACCTTAACAAATTAACCTTTGATCTTTTAAAG 60267
Oy 3123 attcttaaaaaaacatctgtatctatcatatcatatgagctcttgagagactgacttccata 3182
Db 60266 ATTATTAACCACTACGCTGTATGATATGACTCTTCAAGACCCACAGCTTTCACTATAG 60207
Oy 3183 ctacagagatgctcctcatgctgtagcctatataaactgacagagatgatttcc-agagctgc 3241
Db 60206 CTTAGAGATATTTCTCTCTGTGTAGTACATATGACATTTGCAAGTTAAATTTCTTGAGTGC 60147
Oy 3242 tcgatactgtaatactacac 3261
Db 60146 TTTATCTCTTATTTTCTC 60127

RESULT 14
LOCUS G30939 385 bp DNA STS 29-SEP-1998
DEFINITION DRES50 Human DRES Homo sapiens STS genomic clone 22368 3', sequence
tagged site.
ACCESSION G30939
VERSION G30939.1 GI:1724032
KEYWORDS STS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 385)
Bafli, S., Borsani, G., Rossi, E., Bernard, L., Guffanti, A.,
Rubboli, F., Marchitelli, A., Giglio, S., Coluccia, E., Zollo, M.,
Zuffardi, O. and Ballabio, A.
Identification and mapping human cDNAs homologous to Drosophila
mutant genes through EST database searching
Nat. Genet. 13 (2), 167-174 (1996)
96225443

COMMENT
Contact: Ballabio Andrea
Telethon Institute of Genetics and Medicine, TIGEM
Telethon Institute of Genetics and Medicine, TIGEM
via Olgettina 58, 20132 Milano, Italy
Tel: +39-2-215601
Fax: +39-2-21560220
Email: dres@tigem.it
Primer A: GTAAACAATGAATGACGACG
Primer B: GGCGTGAAGAATGACCTAC
STS size: 157
PCR Profile:
Presoak: 94 degrees C for 2 minutes
Denaturation: 94 degrees C for 45 seconds
Annealing: 57 degrees C for 45 seconds
Polymerization: 72 degrees C for 45 seconds
PCR Cycles: 35
Thermal Cycler: MJR PTC-200
Protocol:
Template: 25 ng
Primer: each 1 uM
MgCl2: each 220 uM
Taq polymerase: 0.025 units/ul
JOURNAL
TITLE Homo sapiens chromosome 15, clone RP11-746f15
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
REFERENCE 2 (bases 1 to 160979)

```

```

Buffer: MgCl2: 1.5 mM
        KCl: 50 mM
        Tris-HCl: 10 mM
        pH: 8.3

The 5' end of the cDNA clone 22368 is homologous to the staufen
Drosophila gene product. EST from which STS is derived: T89105.
Forward primer is 70-1233; Reverse primer is 70-1234.
Location/Qualifiers
    source
    1. 385
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /map="20, 12.10 cR from GCT10P11 (lod >3.0)"
    /clone="22368"
    /clone_11b="Human DRES"
    STS
    primer_bind
    29. 185
    29. 49
    BASE COUNT 135 a 79 c 71 g 100 t
    ORIGIN

Query Match 10.8%; Score 380.2; DB 75; Length 385;
Best Local Similarity 99.28; Pred. No. 2.1e-78;
Matches 382; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 3069 tcaatttgtaatgcaaacagagacttaaatgaacttgatcctctgtttaagaattatc 3128
Db 385 TTTATTTTGTATGCAAAAGAGACTTAATGACTTGAATCTGTGTTTAAAGATTTAT 326
Oy 3129 aaaaaacatgctgctatcatatcatatgagctcttgagagactgacttccatactacag 3188
Db 325 AAAAAACATTTGTATCTATCATATATGAGCTCTTGAGAGACTTACTCTTCACTACACTACAG 266
Oy 3189 gatagatctccatgctgctcatataaaccctgcagagtgatattccagaatgctcgatc 3248
Db 265 GATATGATCTTCATGATGCTCATATTAACCTGCAGAGTATTTCCAGATGCTCGATAC 206
Oy 3249 tcttaatacatcctcatatgagctgcaaaagaatgacacagcttctgtatcacagctg 3308
Db 205 TGTATTTACATCTCCATTAAGGCGCTGAAAAGAAATGACCTTGTGTATACAGCTGTG 146
Oy 3309 ttgctttagatgctgtgctgctgctacacagaagtgctgacagagctcgtgctgctg 3368
Db 145 TTGCTTTTGTATGTTGTTACTGTACACAGAAAGTGTGCACAGAGCTGCTGCTGCTGTG 86
Oy 3369 ccgtatggaacacctgtagccctgctgagtgtaagtaactgcttccatcatgtttacgct 3428
Db 85 CCGTATGGAAGCCCTGTAGCCCTCGCGAGTTAAGTACGCTTCATTCATTTGTTACGCT 26
Oy 3429 ggaatttctcccatggaatgta 3453
Db 25 GGAATTTTCTCCCATGGAATGTA 1

RESULT 15
AC055879 160979 bp DNA HTG 22-MAY-2000
LOCUS AC055879
DEFINITION Homo sapiens chromosome 15 clone RP11-746f15 map 15, WORKING DRAFT
SEQUENCE, 25 unordered pieces.
ACCESSION AC055879
VERSION AC055879.2 GI:8017940
KEYWORDS HTG; HTGS_PHASEL; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 (bases 1 to 160979)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 15, clone RP11-746f15
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 160979)

```


RESULT 13	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AL158850/c	192597 bp	Homo sapiens chromosome 6 clone RP3-522B19, ***	AL158850	AL158850.4	HTG; HTGS_PHASE1; HTGS_DRAFT.	human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	Submitted (05-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquere@sanger.ac.uk	Direct Submission	On Aug 13, 2000 this sequence version replaced gi:9212635.

```

Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: hummery@sanger.ac.uk
----- Project Information
Center project name: d5522819
----- Summary Statistics
Assembly program: XGAP, version 4.5
Sequencing vector: M13, M7815; 18% of reads
Sequencing vector: plasmid; L08753; 8% of reads
Chemistry: Dye-terminator Big Dye; 7% of reads
Chemistry: Dye-primer-amesham; 2% of reads
Consensus quality: 185525 bases at least Q40
Consensus quality: 188833 bases at least Q30
Consensus quality: 190005 bases at least Q20
Insert size: 191597; sum-of-contigs
Insert size: 14416; 9.6% agarose-fp
Quality coverage: 5.71x in Q20 bases; sum-of-contigs
Quality

```

```

coverage: 8.21x in Q20 bases; agarose-ip
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 13837: contig of 13837 bp in length
*
* 13838 13937: gap of 100 bp
*
* 13938 17476: contig of 3539 bp in length
*
* 17477 17576: gap of 100 bp
*
* 17577 20005: contig of 2429 bp in length
*
* 20006 20105: gap of 100 bp
*
* 20106 32070: contig of 11965 bp in length
*
* 32071 32170: gap of 100 bp
*
* 32171 39772: contig of 7602 bp in length
*
* 39773 39872: gap of 100 bp
*
* 39873 43471: contig of 3599 bp in length
*
* 43472 43571: gap of 100 bp
*
* 43572 51795: contig of 8224 bp in length
*
* 51796 51895: gap of 100 bp
*
* 51896 178838: contig of 126543 bp in length
*
* 178839 178938: gap of 100 bp
*
* 178939 181919: contig of 2981 bp in length
*
* 181920 182019: gap of 100 bp
*
* 182020 184231: contig of 2212 bp in length
*
* 184232 184331: gap of 100 bp
*
* 184332 192597: contig of 8266 bp in length.
*
location/qualifiers
1..192597

```

misc_feature	20106..32070	/note="assembly-fragment:03060 fragment_chain:2"
misc_feature	32171..39772	/note="assembly-fragment:00945 fragment_chain:2"
misc_feature	39873..43471	/note="assembly-fragment:00013 fragment_chain:2"
misc_feature	43572..51795	/note="assembly-fragment:02225 fragment_chain:3"
misc_feature	51896..178838	/note="assembly-fragment:01767 fragment_chain:3"
misc_feature	178839..181919	/note="assembly-fragment:01637"
misc_feature	182020..184231	/note="assembly-fragment:02408"
misc_feature	184332..192597	/note="assembly-fragment:02960"
BASE COUNT	58324 a 39515 c 38313 g 55431 t 1014 others	
ORIGIN		